North Carolina Digital Preservation Policy

April 2014

State Archives of North Carolina
NATURAL AND CULTURAL RESOURCES

Contents

Executive Summary

The North Carolina Digital Preservation Policy (Policy) governs the operation, management, and scope of the North Carolina digital preservation program (program), established by the North Carolina Department of Cultural Resources in accordance with state law that authorizes the State Archives of North Carolina and State Library of North Carolina to collect and preserve state government information for public access. The program, which supports the preservation and access of authentic digital records throughout their entire lifecycle, provides for strategic decisions about technology and workflows as well as formal structures whereby those records are made available to citizens as part of North Carolina's heritage. Operating principles – from robust administrative support to adoption of best practices and accepted preservation standards – guide the program and its stakeholders. These operating principles call for collaborative efforts stewarded by the State Archives and State Library, efforts that involve managing a record throughout its lifecycle in cooperation with participants from all parts of state government.

Purpose

This policy governs the operation, management, and scope of the North Carolina digital preservation program, established by the North Carolina Department of Cultural Resources in accordance with state law that authorizes the State Archives of North Carolina and State Library of North Carolina to collect and preserve state government information for public access. This policy allows the State Archives and State Library to ensure the historical and cultural legacy of the state manifested in digital form will not be lost to future generations. This commitment is guaranteed through the adoption of strategic and technical measures for the long-term preservation of North Carolina's digital state government assets, including digital content produced by state and local governments, contained within private collections, and produced by the staff of the State Archives and the State Library. The adoption of the following policies to capture, preserve, and make accessible the digital record of North Carolina's government and its citizens will ensure the administrative integrity, information authenticity, government transparency, and legislative compliance of the State's digital assets. This policy outlines the State Archives' and State Library's strategic approach to minimizing the risks associated with technological changes and ensuring that the records remain reliable and accessible over time.

This document is published in connection with the *DCR Digital Preservation Policy Framework*,[1] formally adopted in 2005 by the State Archives of North Carolina, the State Library of North Carolina, and the executive management of the Department of Cultural Resources. It is consistent with the State Archives' core mandates as authorized by GS §132 and GS §121 to "preserve and administer, in the North Carolina State Archives, such public records as may be accepted into its custody, and to collect, preserve, and administer private and unofficial historical records and other documentary materials relating to the history of North Carolina…"[2] and the State Library's mandate GS § 125, which charges the State Library as the "official, complete, and permanent depository for all State publications."[3]

This policy acknowledges that a viable and successful digital preservation program requires an ongoing commitment of resources—financial, technical, and human—from the State Archives and State Library. As technologies mature and become viable options for digital preservation, State Archives and State Library staff will identify, assess, procure, and implement a technical digital preservation environment solution according to the timeline and goals set forth in the *North Carolina Department of Cultural Resources Digital Preservation Plan: 2013-2015*.

Audience

The intended audience for this policy includes staff of the Department of Cultural Resources, government employees from both state agencies and local governments, as well as donors of private papers and collections whose digital assets the State Archives and State Library preserve in their collections.

Policy Updates

This policy may be revised and updated at any time, in order to reflect technological, infrastructural, and operational developments in the digital preservation program. The policy is intended to be updated at least biennially.

Objectives

The digital preservation program is intended to preserve, maintain, and make accessible digital state government assets as well as assets in the private collections, including those originating in digital format, sometimes referred to as "born digital", and those transformed to a digital form and referred to as "digitized."

---

[1] North Carolina Department of Cultural Resources. Digital Preservation Policy Framework, September, 2005. http://digitalpreservation.ncdcr.gov/asgii/publications/policy_framework.pdf

[2] G.S. §121-4 "Powers and duties of the Department of Cultural Resources"

[3] G.S. §125-11.7 "State Library designated the official depository for all State publications"

The primary objectives of the digital preservation program are to:

- Preserve digital North Carolina assets that are either born digital or digitized. Retain a copy of the original bit stream of born digital material for authenticity purposes, as well as potential access in the future. These assets include state government information as well as materials in its private collections.

- Where necessary, migrate records into formats that are acceptable for long-term preservation and access.

- Make strategic decisions, based on long-term resources, system sustainability, and identified efficiencies, to establish processes and technologies that support both the Trusted Digital Repository standard (ISO 16363) and the Open Archival Information System (OAIS) reference model (ISO 14721).

- Ensure record authenticity through preservation and technical solutions that promote and maintain composition (especially for assets with multiple objects or files), lineage, fixity, and validity.

- Establish operational processes and procedures to meet archival requirements pertaining to provenance, chain of custody, authenticity, and integrity.

- Provide tools that facilitate the discovery of online digital assets from the State Archives and State Library.

- Define policies and procedures for the preservation and availability of digital assets while maintaining intellectual property ownership and rights.

- Make strategic decisions, based on long-term resources, system sustainability, and identified efficiencies, to maintain a technical and operational infrastructure based on the conceptual framework of ISO 14721 Open Archival Information System (OAIS). Automate preservation processes as technical capabilities and resources— financial and human—allow.

- Support and manage storage media for archival copies (preservation master, security copies, and access copies) in accordance with environmental, quality control, security, and other standards. Sustain a storage architecture that provides strategically selected redundancy in the event of technical failure, natural disaster, or human error.

Scope

The digital preservation program is responsible for identifying, securing, and providing the means to preserve and ensure ongoing access to digital state government assets. Digital assets are those electronic objects that have been identified as having enduring cultural, historical, informational, and/or evidentiary value to the State of North Carolina. Examples of digital assets include relatively simple formats such as word processing documents, spreadsheets, digital publications, or digital images, and complex application-specific digital assets such as email, websites, databases, and geospatial datasets. To become part of digital preservation program, a digital asset will be evaluated in accordance with records retention and disposition schedules and appraisal of its long-term or permanent value to the State of North Carolina. Digital assets selected for permanent retention will have ongoing usefulness or significance, based on the administrative, legal, fiscal, evidential, or historical information they contain that justifies their continued preservation. The following types of records will be included under this program:

- Records and donations that contain scanned or digitized material for which no analog counterpart exists.

- Born-digital assets produced by state government agencies

- Born-digital assets produced by local government agencies

- Materials that are part of a private collection donated to the State Archives

- Archives or Library materials that are digitized and identified for long-term digital retention.

With the ever growing volume of digital information, the State Archives and the State Library may need to allocate priorities for preservation action based on the relative significance of digital assets and the technical complexity of preserving and ensuring access to those assets.

Incentives and Challenges

The State Archives and State Library have a legal mandate to preserve the record of North Carolina's government and its citizens. This mandate has profound implications for the efficient management of democratic government and cultural heritage of the people of North Carolina. Though the costs of preserving digital assets may be high, the costs and implications of failing to preserve them are even higher. Loss of a digital asset not only means loss of the state's historical record, but also loss of the original investment of staff time and resources required to create the asset. This is especially true for born-digital materials, which can be lost either at the point where the asset becomes technically impossible to access or where access is so cost prohibitive (time, money, software) that recovery of the asset is infeasible. Without the adoption of strategic and tactical measures to deploy adequate resources to capture, preserve, and make accessible the digital record of North Carolina's government and its citizens, the historical and cultural legacy of the state manifested in digital form will be lost not only to future generations, but to today's generation.

The primary challenge is the technical complexity inherent in the design, development, and operation of a digital preservation system. Archivists and librarians are in a perpetual race against the technology clock as hardware, software, and media formats evolve and become obsolete. Hardware, software, and network failures threaten to corrupt digital assets. Natural disasters can destroy data centers. Human failure is an ever present challenge as digital assets are often unintentionally modified or deleted. There are also operational challenges such as adequately describing digital assets so that they are accessible, ensuring their technical characteristics are captured, and their archival provenance is maintained. Users must be able to access electronic records. Finding an appropriate user interface or building and supporting one is critical to enabling users to search and locate, and optimally retrieve digital assets that are of interest.

Operating Principles

The State Archives' and State Library's digital preservation objectives are best achieved through the development of practices that comply with an adequate, coherent, and widely understood framework for reliable, accountable, and manageable digital archives. The digital preservation program is based on the following principles:

- A collaborative approach to realizing a digital preservation program involving stakeholders, both within North Carolina and the nation, at all levels of planning, development, and implementation.

- Ongoing, sustained support and engagement from the Department of Cultural Resources (DCR) management, information technology professionals, and the content and records creators.

- Development of an interoperable digital repository framework using the best technology available— commercially or in house—incorporating open source options whenever feasible.

- Implementation of a range of community-based standards, best practices, and national and/or international standards that inform preservation procedures and technology as well as archival requirements such as provenance, chain of custody, intellectual property rights, and authenticity.

- To the best our capability, compliance with ISO 14721 the Open Archival Information System reference model and ISO 16363 Trusted Digital Repository Standard for preservation as well as with all requirements of the North Carolina Uniform Electronic Transmissions Act (UETA) that pertain to the delivery of digital content to the public .

- Development of consistent documented guidelines and procedures for each stage of the lifecycle (i.e., creation, selection, acquisition, ingest, preservation action—including reformatting and producing derivatives—storage, identification/cataloging, access and use, transform, dispose).[4]

- Ongoing communication with peer institutions about opportunities for improved processes, monitoring the landscape for services offered by third party vendors and participation in national research projects regarding digital preservation.

- Review of policies and procedures on a regular basis, taking into account changes in the organizational, legal, and technical environment of the State Archives, State Library, and government statewide.

Roles & Responsibilities

The Department of Cultural Resources (DCR) serves as the lead agency for digital preservation in North Carolina state government responsible for supporting and sustaining the statutory responsibilities of the State Archives and the State Library to preserve and provide public access to print and digital state government information and materials in private collections and coordinating stakeholder involvement and collaboration within and outside state government.   Primary stakeholders in the digital preservation of state government information include:

- **DCR management**: It is the responsibility of DCR management to provide adequate managerial and financial commitment for the digital preservation program. As resource commitments are made, management should take into consideration the extent to which digital preservation can enhance or replace existing services and functions.

- **State Archives and State Library Staff**: Traditionally, responsibility has been divided between the State Archives and the State Library for records and publications. Expertise from both programs will ensure that *all* digital content in scope is collected and preserved.

- **Digital record producers / contributors:** The producers and contributors of the digital records will be responsible for complying with established submission requirements and working with the management of the digital archive to ensure successful transfer. Records producers will have the same financial responsibilities for preparing digital records as they currently do analog records.

- **State agencies and local government**: As with analog materials, content creators help ensure that digital content is properly identified and collected by DCR. However, because of the immensely greater risk of loss associated with digital formats, there will need to be greater collaboration between creators and DCR to manage digital assets throughout their entire lifecycle.

- **DCR-IT**: The preservation of digital materials requires a much higher level of involvement of information technology experts than needed with analog materials. DCR-IT needs to assist the State Archives and State Library with the purchase, configuration, customization (custom development), and maintenance of hardware and software solutions as well as the network infrastructure.

- **Information Technology Services (ITS)**: For the program to have a true lifecycle approach to the management of digital assets, recognition and collaboration with statewide ITS management is needed to manage content from creation to disposition.

---

[4] Lifecycle management of digital assets is based on the Digital Curation Centre's Lifecycle Model.

- **Citizens of North Carolina**: The citizens of North Carolina have the right to access essential government information in a timely and efficient fashion.

Collaboration / Cooperation

The Department of Cultural Resources is committed to supporting collaboration internally between the State Archives and the State Library as well as externally with local and state government and other institutions, both within the state of North Carolina and nationally to advance the development of the digital preservation program, share lessons learned with other digital preservation programs, extend the breadth of available expertise, and extend the digital content that is available to the public through cooperative efforts.

 The Department of Cultural Resources will:

- Identify appropriate partners and stakeholders able to contribute to the statewide effort.

- Establish agreements regarding responsibilities and roles.

- Pursue agreements that provide a reliable basis for ongoing accessibility over time.

- Help identify and develop policies, procedures, and tools to support the management and preservation of digital information.

- Work with creators, publishers, and re-users of digital content to encourage practices that will enable, rather than hinder, preservation.

- Work with government agencies and officials to develop legislative and funding frameworks that will enable cost-effective preservation.

Technical Principles

Authenticity

The State Archives and State Library strive to ensure the authenticity of digital resources; the mutable nature of digital assets opens the possibility for unauthorized and undetectable changes. Confidence in the authenticity of digital records over time is particularly crucial owing to the ease with which alterations can be made. From the moment that digital resources are acquired, the State Archives and State Library undertake protective procedures to prevent, discover, and correct loss or corruption of digital assets due to either inadvertent or malicious intent. In addition, the Archives and Library staff will endeavor to secure supporting evidence, ideally in the form of metadata, from those creating the resources. This will enable those who access the resources to evaluate the authenticity of all preserved digital resources.

Metadata

Metadata (descriptive, preservation and administrative) is fundamental to preserving and providing access to the State Archives' and State Library's digital assets. The preservation process includes maintenance of metadata submitted with the digital asset as well as the creation of additional metadata to manage the long-term/active preservation and access of that asset. The Archives and Library staff are committed to identifying and collecting the needed metadata to preserve digital assets in its collection and to provide access to those assets.

Formats

In acquiring digital materials, the State Archives and State Library will define and communicate levels of preservation appropriate to each type of format. The State Archives and State Library reserve the right to assign different

preservation standards to different file formats. Please see *File Formats for Transfer of Electronic Records to the State Archives of North Carolina* document (**www.ncdcr.gov/archives**) for current standards.

Access/Use

The purpose of preserving digital assets is to ensure that the assets remain accessible in the future. The State Archives and State Library must provide access using the most up-to-date technology available. For data that end users may download to their own computers, users will need to be informed of what software is necessary to render the digital record. Where possible, the State Archives and State Library will strive to make digital records available in openly-documented formats for which rendering tools are readily available (e.g. PDF files, TIFF files, WAVE files).

In some cases, the digital record may be so tightly coupled with its originating software program that the record can only be rendered through the application used to create the digital record (e.g. Esri Geodatabases). In order to ensure continued access, the State Archives and State Library may migrate assets to a new format. Some information may be lost during the migration but the fundamental information in the records will be accessible. However, the State Archives and State Library, as part of their preservation program, will maintain the original bit stream and make that available should it be needed. Sensitive and confidential information will require appropriate restrictions for access and use. Content intended for public dissemination will be available for public use.

References

Consultative Committee for Space Data Systems. "Reference Model for an Open Archival Information System (OAIS), CCSDS 650.0-M-2, Magenta Book." June 2012. Retrieved 7/24/2013 from: **http://public.ccsds.org/publications/archive/650x0m2.pdf**

Cornell University Library. "Cornell University Library Digital Preservation Policy Framework." December 2004. Retrieved 5/11/2011 from: **http://commondepository.library.cornell.edu/cul-dp-framework.pdf**

Digital Curation Centre. "DCC Curation Lifecycle Model." Retrieved 5/6/2011 from: **http://www.dcc.ac.uk/resources/curation-lifecycle-model**

Georgia Archives. "Digital Preservation Policy." Retrieved 5/11/2011 from: **http://www.sos.ga.gov/archives/who_are_we/rims/digital_History/policies/policy%20-%20Digital%20Preservation%20Policy.pdf**

OCLC. Trustworthy Repositories Audit & Certification: Criteria and Checklist. 2007. Retrieved 5/11/2011 from: **http://www.crl.edu/sites/default/files/attachments/pages/trac_0.pdf**

Pearce-Moses, Richard. "A Glossary of Archival and Records Terminology." Society of American Archivists, 2005. Retrieved 5/12/2011 from: **http://www.archivists.org/glossary/**

RLG-OCLC Working Group on Digital Archive Attributes. "Trusted Digital Repositories: Attributes and Responsibilities." Mountain View, CA: RLG, May, 2002. Retrieved 5/12/2011 from: **http://www.oclc.org/research/activities/past/rlg/trustedrep/repositories.pdf**

Yale University Library. "Yale University Library Digital Preservation Policy." Retrieved 5/11/2011 from: **http://www.library.yale.edu/iac/DPC/final1.html**

Glossary

Archival Information Package (AIP): An (OAIS) Information Package consisting of the Content Information and the associated Preservation Description Information which is preserved within an Open Archival Information System.

Access: The permission to locate and retrieve information for use within legally established restrictions of privacy, confidentiality, and security clearance.

Authenticity: A quality of a digital resource to be judged trustworthy and genuine based on internal and external evidence. (Yale)

Born-digital: Information created in electronic format. Born-digital information is distinguished from information that was created in another format and digitized through scanning or digital photography, creating a digital surrogate. A document created using a word processor may be described as born-digital.

Data Object (OAIS): Either a Physical Object or a Digital Object.

Digital asset: digital content that has value based on the administrative, legal, fiscal, evidential, or historical information it contains. Digital assets can consist of born-digital or digitized content.

Digital (electronic) content: Data or information that has been captured and fixed for storage and manipulation in an automated system and that requires the use of the system to render it intelligible by a person. This definition encompasses records and publications.

Digitization: The process of transforming analog material into digital (electronic) form, especially for storage and use in a computer. Also referred to as scanning or imaging.

Digital Preservation: The managed activities necessary for ensuring both the long-term maintenance of a byte stream and continued accessibility of its contents. [RLG-OCLC: "Trusted Digital Repositories: Attributes and Responsibilities"]

Dissemination Information Package (DIP): The (OAIS) Information Package derived from one or more Archival Information Packages (AIPs) received by the Consumer (end user) in response to a request to the Open Archival Information System.

Fixity: The quality of being stable and resisting change. Mechanisms such as checksums or cyclical redundancy checks can record a fixity value for a digital object, and can be used as a basis of later comparison to ensure that the digital object has not been altered.

Fixity Information: The information which documents the authentication mechanisms and provides authentication keys to ensure that the Content Information object has not been altered in an undocumented manner. An example is a checksum or a Cyclical Redundancy Check code for a file.

Information Object (OAIS): A Data Object together with its Representation Information.

Metadata: Information that describes significant aspects of a resource. Preservation metadata are required to describe, manage, and preserve digital resources over time and will assist in ensuring essential contextual, administrative, historical, and technical information are preserved along with the digital resource. (Yale)

Life Cycle: The framework for understanding the sequential or cyclical sequence of activities that digital assets undergo during their existence. (Yale)

Open Archival Information System (OAIS): An archive, consisting of an organization of people and systems that has accepted the responsibility to preserve information and make it available for a Designated Community. It meets a set of

responsibilities, as defined in 3.1, that allows an OAIS archive to be distinguished from other uses of the term archive. The term Open in OAIS is used to imply that this Recommendation and future related Recommendations and standards are developed in open forums, and it does not imply that access to the archive is unrestricted.

Permanent value: The ongoing usefulness or significance of records, based on the administrative, legal, fiscal, evidential, or historical information they contain that justifies their continued preservation.

Preservation (and digital preservation): The process of protecting materials from deterioration or damage. Digital preservation typically centers on the choice of interim storage media, the life expectancy of a digital imaging system, and the expectation to migrate the digital files to future systems while maintaining both the full functionality and the integrity of the original digital system.

Preservation: A series of active managed activities and processes involved in the physical and intellectual protection of administrative, legal, fiscal, evidential, historical information and cultural materials. Preservation encompasses a host of policies, procedures, and processes that together sustain access or prevent further deterioration to the materials chosen to save. (Yale)

Provenance: The source and ownership history of a (digital) resource. (Yale)

Representation Information (OAIS): The information that maps a Data Object into more meaningful concepts. An example is the ASCII definition that describes how a sequence of bits (i.e. a Data Object) is mapped into a symbol.

Repository: An organization that intends to maintain information for access and use.

Submission Information Package (SIP): An (OAIS) Information Package that is delivered by the Producer/Contributor for use in the construction of one or more Archival Information Packages (AIPs).

Trusted Digital Repository: A repository whose mission is to provide reliable, long-term access to managed digital resources to its designated community, now and in the future. A critical component is the ability to prove reliability and trustworthiness over time. [RLG-OCLC: "Trusted Digital Repositories: Attributes and Responsibilities"]

Appendix A: Open Archival Information System (OAIS) Model
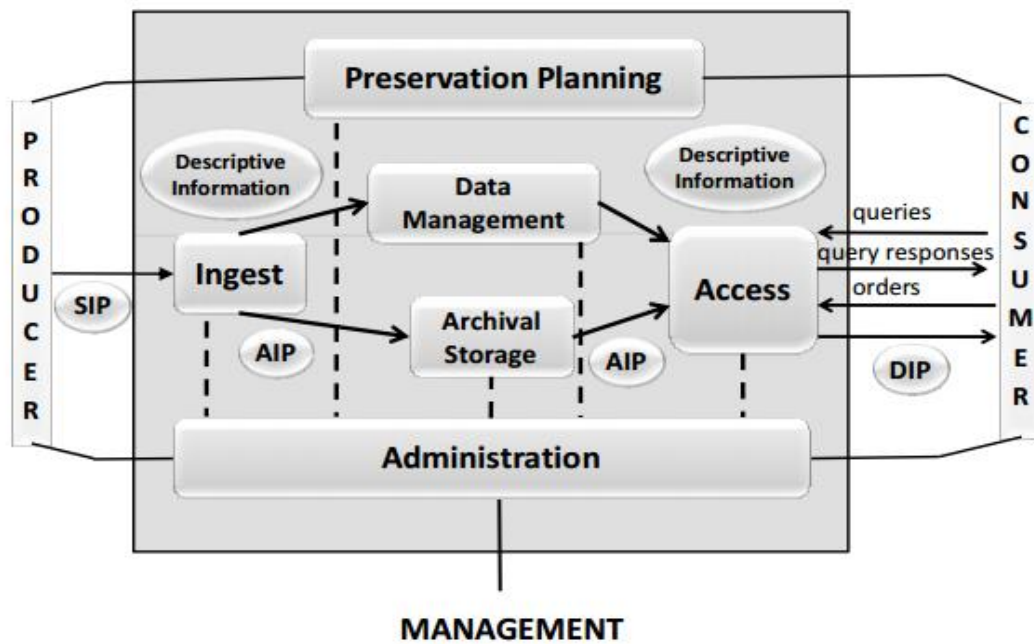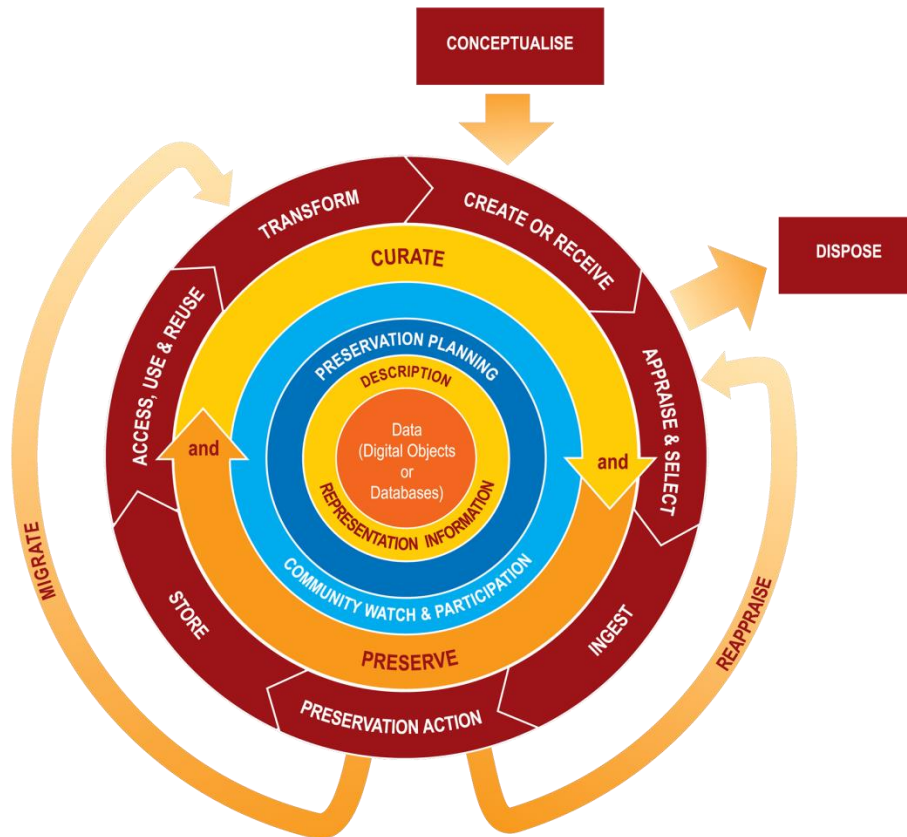

MANAGEMENT

Figure 1. Consultative Committee for Space Data Systems. "Reference Model for an Open Archival Information System (OAIS), CCSDS 650.0-M-2, Magenta Book." June 2012. Retrieved 7/24/2013 from: **http://public.ccsds.org/publications/archive/650x0m2.pdf**.

Appendix B: Digital Curation Lifecycle Model

The Curation Lifecycle

The DCC Curation Lifecycle Model provides a graphical high level overview of the stages required for successful curation and preservation of data from initial conceptualisation or receipt. The model can be used to plan activities within an organisation or consortium to ensure that all necessary stages are undertaken, each in the correct sequence. The model enables granular functionality to be mapped against it; to define roles and responsibilities, and build a framework of standards and technologies to implement. It can help with the process of identifying additional steps which may be required, or actions which are not required by certain situations or disciplines, and ensuring that processes and policies are adequately documented.

Data (Digital Objects or Databases)

Data, any information in binary digital form, is at the centre of the Curation Lifecycle. This includes:

Digital Objects	- Simple Digital Objects are discrete digital items; such as textual files, images or sound files, along with their related identifiers and metadata. - Complex Digital Objects are discrete digital objects, made by combining a number of other digital objects, such as websites.
Databases	Structured collections of records or data stored in a computer system.

Full Lifecycle Actions

Description and Representation Information	Assign administrative, descriptive, technical, structural and preservation metadata, using appropriate standards, to ensure adequate description and control over the long-term. Collect and assign representation information required to understand and render both the digital material and the associated metadata.
Preservation Planning	Plan for preservation throughout the curation lifecycle of digital material. This would include plans for management and administration of all curation lifecycle actions.
Community Watch and Participation	Maintain a watch on appropriate community activities, and participate in the development of shared standards, tools and suitable software.
Curate and Preserve	Be aware of, and undertake management and administrative actions planned to promote curation and preservation throughout the curation lifecycle.

Sequential Actions

Conceptualise	Conceive and plan the creation of data, including capture method and storage options.
Create or Receive	Create data including administrative, descriptive, structural and technical metadata. Preservation metadata may also be added at the time of creation. Receive data, in accordance with documented collecting policies, from data creators, other archives, repositories or data centres, and if required assign appropriate metadata.
Appraise and Select	Evaluate data and select for long-term curation and preservation. Adhere to documented guidance, policies or legal requirements.
Ingest	Transfer data to an archive, repository, data centre or other custodian. Adhere to documented guidance, policies or legal requirements.
Preservation Action	Undertake actions to ensure long-term preservation and retention of the authoritative nature of data. Preservation actions should ensure that data remains authentic, reliable and usable while maintaining its integrity. Actions include data cleaning, validation, assigning preservation metadata, assigning representation information and ensuring acceptable data structures or file formats.
Store	Store the data in a secure manner adhering to relevant standards.
Access, Use and Reuse	Ensure that data is accessible to both designated users and reusers, on a day-to-day basis. This may be in the form of publicly available published information. Robust access controls and authentication procedures may be applicable.
Transform	Create new data from the original, for example - By migration into a different format. - By creating a subset, by selection or query, to create newly derived results, perhaps for publication.

Occasional Actions

Dispose	Dispose of data, which has not been selected for long-term curation and preservation in accordance with documented policies, guidance or legal requirements. Typically data may be transferred to another archive, repository, data centre or other custodian. In some instances data is destroyed. The data's nature may, for legal reasons, necessitate secure destruction.
Reappraise	Return data which fails validation procedures for further appraisal and reselection.
Migrate	Migrate data to a different format. This may be done to accord with the storage environment or to ensure the data's immunity from hardware or software obsolescence.

Figure 2. Digital Curation Center, "Digital Curation Lifecycle Model." Retrieved 7/24/2013 from:
http://www.dcc.ac.uk/sites/default/files/documents/publications/DCCLifecycle.pdf